EXHIBIT I

The Value of Quantitative Electroencephalography in Clinical Psychiatry: A Report by the Committee on Research of the American Neuropsychiatric Association

Kerry L. Coburn, Ph.D.
Edward C. Lauterbach, M.D.
Nash N. Boutros, M.D.
Kevin J. Black, M.D.
David B. Arciniegas, M.D.
C. Edward Coffey, M.D.

The authors evaluate quantitative electroencephalography (qEEG) as a laboratory test in clinical psychiatry and describe specific techniques, including visual analysis, spectral analysis, univariate comparisons to normative healthy databases, multivariate comparisons to normative healthy and clinical databases, and advanced techniques that hold clinical promise. Controversial aspects of each technique are discussed, as are broader areas of criticism, such as commercial interests and standards of evidence. The published literature is selectively reviewed, and qEEG's applicability is assessed for disorders of childhood (learning and attentional disorders), dementia, mood disorders, anxiety, panic, obsessive-compulsive disorder, and schizophrenia. Emphasis is placed primarily on studies that use qEEG to aid in clinical diagnosis, and secondarily on studies that use qEEG to predict medication response or clinical course. Methodological problems are highlighted, the availability of large databases is discussed, and specific recommendations are made for further research and development. As a clinical laboratory test, qEEG's cautious use is recommended in attentional and learning disabilities of childhood, and in mood and dementing disorders of adulthood.

(The Journal of Neuropsychiatry and Clinical Neurosciences 2006; 18:460–500)

Received and accepted March 3, 2006. Dr. Coburn is affiliated with the Department of Psychiatry and Behavioral Sciences, Mercer University School of Medicine, Macon, Georgia. Dr. Lauterbach is affiliated with the Division of Adult and Geriatric Psychiatry, Mercer University School of Medicine, Macon, Georgia. Dr. Boutros is affiliated with the Department of Psychiatry and Neurology, Wayne State University, School of Medicine, Detroit, Michigan. Dr. Black is affiliated with the Department of Psychiatry, Neurology, and Radiology, Washington University School of Medicine, St. Louis, Missouri. Dr. Arciniegas is affiliated with the Department of Psychiatry and Neurology, University of Colorado School of Medicine, Denver, Colorado. Dr. Coffey is affiliated with the Henry Ford Health System, Detroit, Michigan. Address correspondence to Dr. Coburn, 655 First St., Macon, GA 31201; Coburn_kl@Mercer.edu (E-mail).

Quantitative EEG (qEEG) involves computer-assisted imaging and statistical analysis of the EEG for detecting abnormalities, assisting the physician in making a diagnosis, and other purposes relating to patient care. Among the techniques of functional brain imaging, qEEG offers many advantages. It has an ideal temporal resolution in the millisecond time domain characteristic of neuronal information processing, employs no ionizing radiation, noninvasively images both excitatory and inhibitory cortical neuronal activity rather than secondary hemodynamic processes, and is relatively inexpensive and portable. Its formerly poor spatial resolution has increased dramatically as channel capacity has expanded from 20 a decade ago to 256 presently, with a 512-channel system expected for commercial release within the next year. Perhaps most importantly, several large qEEG normative (i.e., statistically representative) databases directly relevant to clinical psychiatry are available, and qEEG technology has advanced to the point where two systems have attained FDA approval.

Previous reviews of this area have lumped together two types of studies: those focusing on the direct clinical applicability of currently available qEEG systems and those involving more speculative areas of qEEG research. Consequently, it remains unclear whether qEEG is ready to be used as a standard laboratory test by practicing psychiatrists. A pivotal question remains unanswered concerning the actual clinical utility of qEEG and related electrophysiological methods: are the techniques sufficiently sensitive and specific to answer practical clinical questions about individual patients suffering from recognized psychiatric disorders? This article reviews briefly the types of assessments that comprise the realm of qEEG, the areas of controversy surrounding the techniques, and the published studies applying them to individual patients. The focus of this report is on whether presently available qEEG systems can tell the practicing psychiatrist anything of practical importance about the individual patient sitting across the desk from him. Its conclusions are less glowing than might be expected on the basis of previous reviews because, although qEEG can provide information of direct clinical relevance, even the most sophisticated qEEG systems now available are still very limited. We make specific recommendations regarding qEEG's present clinical utility and areas in which additional research and development are needed.

METHOD

Selection of Literature

The focus of this review on the practical clinical utility of qEEG as a laboratory test in psychiatry requires the exclusion of a vast amount of tangentially related literature. EEG biofeedback ("neurotherapy") is not included because it is a treatment rather than a laboratory test. qEEG drug development studies are excluded because they tend strongly to use group research designs that tell nothing about the individual clinical patient being treated. Psychiatric conditions thought to arise secondary to brain damage (e.g., stroke, traumatic brain injury) or infection (e.g., systemic lupus erythematosus) are excluded, due to the difficulty of determining whether any subsequent psychiatric condition is primary or secondary. Also excluded are psychiatric disorders for which the qEEG literature is sparse, such as Axis II disorders and substance abuse/dependence.

The latter is a particularly messy issue. Substance abuse categories are poorly defined and the criteria are inconsistently applied. Since an alcohol discriminant is included with one qEEG system and has been tested and validated in the literature, it has been included in the Depression section. As for the other recreational drugs, most published studies either involve no predictive classification[1,2] or use psychiatric patients as subjects[3] or, most commonly, fail to control for other drug use or important lifestyle variables. When more and better studies have been completed and when appropriate discriminants are included in qEEG systems, it will be important to review them. But for now a review is premature, particularly since discriminants are not available for use by practicing psychiatrists who may wish to use qEEG. There is an additional aspect of this area that extends it beyond the usual realm of clinical psychiatry. Since most recreational drug use is illegal, and since Thatcher has convinced the courts that (at least for brain injury) qEEG meets admissibility standards, it is especially important to tread carefully through this minefield. Discriminants intended to pick out drug users from a population, where false positives involve legal as well as medical hazards, need to be held to the highest standards of validity and reliability before they are made available for general use.

It was also necessary to omit source localization. LORETA and VARETA (low resolution/variable resolution electromagnetic tomography) are extraordinarily im-

portant advances in qEEG. They can provide unique information regarding the neurophysiological underpinnings of psychiatric disorders, and they bring qEEG squarely into the realm of functional neuroimaging. Unfortunately, these techniques fail the practical utility test. To the practicing clinical psychiatrist it makes no difference whether the major depression in the patient sitting across from him is linked to disturbances in the right prefrontal area or the left prefrontal area. The treatment will be the same. As the field develops, particularly as the current DSM categories are parsed into more meaningful subcategories (by cluster analysis, etc.), it may well be the case that psychiatric disorders linked to abnormalities in specific brain areas will be found to respond to different treatments. Indeed, functional neuroimaging should be at the forefront of neuropsychiatry/behavioral neurology development. But for the time being, LORETA and VARETA are simply irrelevant to the day to day professional life of the average working psychiatrist. For more general reviews of the EEG in psychiatry the reader is directed to the work of Chabot et al.,[4] Boutros,[5] Hughes,[6] Hughes and John,[7] and Small.[8]

The following is not intended to provide a comprehensive review of the qEEG literature. Rather it identifies and discusses selected between-subjects studies that are designed to find either differences between an individual patient and a defined healthy group (for simple EEG abnormality detection) or similarities between an individual patient and a defined clinical group (for diagnostic or other classification). With few exceptions, studies of between-group differences rather than between-subjects differences, and work published in nonpeer-reviewed sources, are excluded. Individual articles in the published literature were located via a literature search of the National Library of Medicine databases using medical subject headings that included {EEG, qEEG, Evoked Potentials, Event-related Potentials} and {Mental Disorders, Psychiatric Disorders, Depression, Schizophrenia, Anxiety Disorder, Mood Disorder, Bipolar Illness}. Additional studies were found in the bibliographies of the located articles. Major qEEG equipment manufacturers and companies offering qEEG services or products were contacted for information.

Diagnostic Terminology

Diagnostic nomenclature has evolved rapidly and can lead to confusion when articles published at different times are compared. A particular hazard involves "rediagnosing" patients in earlier studies by attempting to fit them into current diagnostic categories. For that reason, the authors' original terminology for patient groups has been retained in the discussions below. In contrast, the authors' original terminology for healthy individuals used for control purposes ("normal," "control," "nonpatient," etc.) has been changed to "healthy" or "healthy subjects" for the sake of uniformity.

Types of Assessment

A useful nosology of qEEG and related techniques has been provided by Duffy et al.[9] and Nuwer.[10] Of the two major types of data, the debate has centered on qEEG. Evoked potentials (EPs), event-related potentials (ERPs) and their quantitative counterparts (qEPs and qERPs) have received very little attention. The issues are essentially the same, although clinical qEP/qERP development lags far behind qEEG. The following analysis sequence described by Duffy et al.[9] proceeds from least to most controversial aspects of qEEG.

Visual Analysis Visual analysis of the ink-written EEG by a qualified electroencephalographer remains the gold standard and is the first step in any qEEG analysis. Several authors (e.g., Hughes and John[7]) recommend routine visual EEG screening of newly presenting psychiatric patients, particularly if a complete neurological examination is not routinely performed.[5] The use of "paperless" digital EEG (dEEG) allowing modification of display parameters and electrode montages for visual analysis on a computer screen and easy storage of the EEG record in digital form is noncontroversial once certain minimal technical criteria are met. This lack of controversy is somewhat surprising since there appear to be no studies comparing ink-written to paperless EEG to determine optimal standards of screen resolution, presentation rate, or other display variables. These variables may well exert a strong influence on the detection rate of subtle EEG abnormalities. Even worse, the artifactual production of slow activity through the process of "aliasing"[11,12] is possible if high frequency activity in the EEG is sampled at too low a rate. Nevertheless, there is universal agreement that a traditional visual reading of the EEG constitutes an indispensable first step in qEEG analysis.

Spectral Analysis Conversion of the time domain EEG record (voltage plotted against time) to the frequency domain (amplitude or power plotted against frequency)

using the fast Fourier transformation (FFT) has been widely used by researchers since the 1960s but is only now beginning to be employed by clinical EEG laboratories. The use of stand-alone frequency (spectral) analysis without reference to a normative database, as an adjunct to visual analysis, is relatively noncontroversial. However, here, too, there appear to be no studies comparing the clinical utility of the various analytic algorithms in use. Hanning versus sine versus a multitude of other techniques for handling window edge effects, minimum and maximum epoch lengths, and a host of other questions remain largely unaddressed. FFT spectra showing absolute measures will look very different from those showing relative measures, and spectra showing amplitude in microvolts will appear quite different from those showing power in microvolts squared. But no consideration seems to have been given to the possibility that these differing techniques could mislead the physician. The apparent presumption is that spectral analysis using any variant of the technique can call the physician's attention to frequency domain characteristics of the ink-written or dEEG, which may aid in forming a clinical impression of the overall record.

Univariate Comparison to Normative Healthy Databases

Serious controversy begins when qEEG data recorded from a patient are compared statistically with normative databases, on the assumption that clinically significant psychiatric disturbances may be accompanied by statistically significant abnormalities in brain activity. Comparisons using single (univariate) spectral measures of the EEG (or single qEP/qERP amplitude measures) to compute z-scores reflecting the degree of statistical abnormality of the patient's brain activity (e.g., Biologic's Brain Atlas, Nicolet's Brain Electrical Activity Mapping [BEAM] system) tend to be better accepted than the multivariate measures used for patient classification discussed below. But even univariate comparisons raise statistical issues. In order to achieve Gaussianity and avoid statistical bias, some qEEG systems include a log transformation of the FFT data. Also, artifact elimination from the raw data and concerns about the length of artifact-free data required for stable spectral estimates become important considerations at this level of analysis. Since the spectral composition of brain electrical activity changes systematically as a function of normal aging, the more capable qEEG systems use either age-stratified normative databases (e.g., Biologic's Brain Atlas) or age regression (e.g., Neurometric Analysis System) to enhance sensitivity and specificity while avoiding age-related bias. Aside from aging effects, qEEG test-retest stability is remarkably high, even over several years.[13] For practical clinical applications, most head-to-head comparisons of visually analyzed to computer analyzed EEGs[14] find the computer to have the edge for detecting subtle frequency domain abnormalities. Such detection can then alert the clinician that a reevaluation of the EEG is advisable with attention to certain specific features.

The important epistemological difference between this level of qEEG analysis and conventional EEG, or for that matter techniques, such as positron emission tomography (PET) or single photon emission computed tomography (SPECT), is that conventional EEG does not involve quantitative comparisons with normative healthy or patient databases. It is the difference between having a professional opinion informed by a visual impression alone (EEG, PET, SPECT), and a professional opinion informed by a visual impression supplemented by quantitative information (qEEG). Without a reference database, the physician must rely on an impression. As psychiatry moves toward evidence-based medicine, greater reliance may be placed on quantitative analysis, but at the moment the normative databases simply do not exist for other imaging modalities. Univariate abnormality measures have the advantage of being easy to understand. When displayed as statistical probability maps (SPM; sometimes referred to as statistical parametric maps), they are a valuable aid in patient education since brain areas can be made to "light up" in proportion to the abnormality of their activity. They form vivid illustrations of the clinical point that a brain problem underlies a patient's symptoms. This serves to destigmatize psychiatric disorders (the brain is malfunctioning just as any other organ can), bringing them into the realm of "real" medicine, and to motivate compliance with treatment. The patient may not understand theta band slowing over the left posterior parietal lobe, but he can see clearly the bright red area on his brain map.

Error checking is relatively easy since statistically abnormal univariate measures generally will correspond to visible features in the original EEG recording. However, normative databases differ in their composition and quality; a qEEG measure deemed abnormal by comparison with one may be normal when compared with another. Since most normative databases are proprietary products, they are difficult to compare systematically and generally have not had their details published in

the open literature. For all such comparisons of a patient with a healthy control group, it is assumed that patients and controls differ only in the presence of abnormal brain activity underlying the patient's disorder. Unfortunately, many patients do not match the often-stringent selection criteria for the normative healthy group (e.g., no history of neurological or psychiatric disorder, no first degree relatives with such disorders, no hypertension or diabetes, no psychoactive medications, etc.). Due to these selection criteria, controls tend to carry much less overall medical burden than do patients. It must be realized that statistically, such "hyper-healthy" controls are abnormal. Comparing a patient with a hyper-healthy control group involves two confounded components—the difference between the patient and the normative healthy population (i.e., the "street normal" population of average health, but excluding the specific disorder being investigated) and the difference between the normative healthy population and the hyper-healthy subjects. The use of hyper-healthy subjects as opposed to more carefully matched "street normal" controls inflates the type I (false positive) error rate. In many clinical applications maximizing sensitivity at the expense of specificity is defensible on the grounds that it is of overriding importance to avoid missing an abnormality (i.e., making a Type II error) and that false alarms can be weeded out by subsequent evaluations.[15] But there are costs to oversensitive screening. In addition to engendering fear and anxiety over a false positive result, subjecting patients to further diagnostic evaluation entails financial costs and a reasonable chance of additional harm in terms of discomfort, missed work, needle sticks, radiation, IV contrast, etc. Prichep and John[16] make the sensible suggestion that the threshold for clinical concern should be set with regard to the consequences of false negative and false positive results.

Multivariate Comparison to Normative Healthy and Clinical Databases Although its use in clinical psychiatry is controversial, combining several individual (univariate) qEEG measures into a single multivariate measure may allow individual patients to be classified into categories of clinical interest. These often correspond to specific diagnostic categories (for which the classifiers are relatively well developed), but sometimes relate to more tenuously developed categories of medication responsiveness, clinical course, or other dimensions of psychiatric interest. Patient classifications are based on multivariate analysis of linear combinations of qEEG measures (discriminant functions, or "discriminants"), an approach often termed "neurometric" analysis.[16] (For legal purposes the generic term "neurometric" and its variants should be distinguished from the "Neurometric" and Neurometric Analysis System [NAS] trademarks pertaining to a widely used commercial system. For the didactic purposes of this article the distinction is trivial.) This approach extracts a large number of qEEG features and compares them with a reference database. It is assumed that the more statistically unusual the observation, the more likely it is that the underlying brain system is clinically abnormal. Although statistically significant findings are not pathognomonic, they are intended to draw the physician's attention to features of the underlying EEG that may have been overlooked. At its most basic level this multivariate approach offers a broad post-hoc filter for determining whether the patient's EEG is statistically normal or abnormal, much like the univariate approach described above.

Even greater controversy occurs when multivariate methods are extended beyond simple EEG abnormality detection to classify individual patients on a "best fit" basis into specific clinically defined categories. A composite quantitative profile of the individual's EEG can be statistically defined by the particular pattern of z-score values. Patients within a diagnostic category often have distinctive multivariate profiles that are different from those of patients in other diagnostic categories, suggesting that the descriptive symptomatic taxonomy of DSM-IV and ICD-10 may reflect a biological taxonomy of brain abnormalities, which in turn produces a statistical taxonomy of qEEG results. In principle, once the multivariate statistical profiles of different diagnostic categories have been established and validated, they can be used to help diagnose an individual patient on the basis of the similarity of the patient's multivariate qEEG profile to the previously defined profiles of the diagnostic categories. Clinical qEEG proponents are quick to point out that matching a patient's statistical profile to a normative profile most characteristic of a specific disorder is different from using the technique to automate the diagnostic process itself. FDA approval of the Neurometric Analysis System and the NeuroGuide Analysis System (presently the only two approved systems) is for the post-hoc analysis of the EEG, and its developers repeatedly stress the need for a conservative and cautious approach to the interpretation of results.

The unfamiliar nature of multivariate statistical pro-

cedures has led some to consider them "mysterious" and consequently to be distrustful of neurometrics and related approaches. However, the mathematics are standard techniques[17] and are clearly described in the open literature.[16,18–23] Multivariate procedures certainly are easier to understand than the mathematics underlying three-dimensional MRI image construction or, for that matter, the quantum mechanics underlying a simple transistor, though few would consider transistor radios to be mysterious and worthy of distrust. But the multivariate approach has its limitations. qEEG findings are not pathognomonic and are appropriately used only in conjunction with other clinical information rather than as stand-alone diagnostic classifiers. Additionally, due to its foundation in Bayesian statistics, for this type of multivariate comparison to be valid it is necessary to ensure that the patient belongs exclusively to one of a limited number of categories, usually healthy versus a specific disorder, but sometimes one specific disorder versus another specific disorder. Due to their non-zero false positive rates and the limited number of defined clinical categories, it is inappropriate to use these procedures as a general diagnostic screening test. Difficulties have arisen when naïve users have employed the procedures as a diagnostic filter, running a patient's data against all possible diagnostic classifiers. Additionally, multivariate measures of pathology are more difficult for both doctors and patients to understand than their univariate components. They do not map well and therefore are of less use in patient education. They also are more difficult to check for errors since each univariate component of an abnormal multivariate measure need not in itself be abnormal.

Advanced Techniques Holding Clinical Promise qEEG has been reported to do more than simply assist the physician in detecting EEG abnormalities and forming a diagnosis. In a number of instances, qEEG cluster analysis, which groups individuals on the basis of qEEG features without a priori outcome information, has defined subtypes within a single diagnostic class, suggesting that markedly different pathophysiological processes may produce essentially the same clinical symptoms.[4] Sometimes it is found that individuals within different qEEG clusters respond differently to treatment. Two subtypes of attention deficit/learning disabled children have been found, only one of which responds well to methylphenidate.[24] Similarly, Prichep et al.[25] and Hansen et al.[26] have identified two subtypes of obsessive-compulsive disorder (OCD) patients showing differing responses to selective serotonin reuptake inhibitor (SSRI) medications. Although this aspect of qEEG has not been developed sufficiently for clinical application, a physiological method for predicting a patient's response to a medication could have profound value for clinical care, helping to select the medication most likely to benefit the individual patient and thereby shorten unsuccessful medication trials. This is a developing area of qEEG research. Another technique holding clinical promise is LORETA, which back-projects surface recorded qEEG onto a realistic three-dimensional brain model, optionally the patient's own MRI. A LORETA normative database has been described and validated recently[27] and its potential clinical utility has been demonstrated.[28] It remains to be seen whether LORETA can be developed into a useful clinical laboratory test in psychiatry.

Review of the Present Controversy

The great fear seems to be that unsophisticated practitioners will attempt to use the classification ability of multivariate analysis to substitute for, rather than aid in, clinical diagnosis and treatment selection.[10] This fear is not without substance. During the 1980s one commercial vendor aggressively marketed a qEEG instrument incorporating an early version of the Neurometric system as virtually a stand-alone automatic diagnostic test. Marketing targeted psychiatrists, family practitioners, and other medical specialists unsophisticated in the use of clinical EEG. The system's limitations—particularly those related to recording artifacts and the Bayesian structure of allowable comparisons—were ignored, and there was a very real possibility of harming patients by misinterpretation of the results. Experienced electroencephalographers of the neurological community were quick to voice their concerns[29] and have had a continuing chilling effect on qEEG.

In a 1994 paper on behalf of the American Medical EEG Association, Duffy et al.[9] assessed qEEG's clinical efficacy and set minimum standards for its use. Central to these standards is the requirement that only specifically trained individuals should use this technology. Nuwer,[10] writing for the American Academy of Neurology and the American Clinical Neurophysiology Society, dismissed Duffy's paper out of hand and damned the technique by faint praise. Replies by Hoffman et al.[30] representing the Association for Applied Psychophysiology and Biofeedback and the Society for the Study of Neuronal Regulation, and by Thatcher et al.[31] repre-

senting the EEG and Clinical Neuroscience Society, pointed to bias and sloppy scholarship in the Nuwer report. In 1999, a Texas court held that Nuwer's criticisms of qEEG failed to meet acceptable scientific standards.[32] Chabot et al.[4] and Hughes and John[7] have provided more complete reviews of the qEEG literature while other authors[33,34] have addressed conceptual issues.

Group v. Individual Differences Two basic approaches may be discerned to the study of psychiatric illnesses. One approach compares groups of patients to groups of healthy subjects employing research designs intended to find between-group differences attributable to the illness. An enormous research literature documents significant statistical differences between psychiatric patient groups and healthy control groups on a wide variety of qEEG measures. Such between-group designs yield a great deal of information about the workings of the normal brain and the functional alterations characteristic of psychiatric disorders. Examples include studies of the qEEG in schizophrenia,[35] dementia,[36] and depression.[37] The practical clinical problem is that even very significant between-group statistical differences on a measure do not necessarily mean that the measure is capable of classifying individuals into their respective groups with any useful degree of accuracy.[38,39] Unfortunately, much of the literature cited in support of clinical qEEG[4,7] is made up of papers, such as these—good science with unclear clinical application.

A second approach focuses on the individual patient, using qEEG measures to detect abnormalities broadly, and more narrowly to help classify the patient into a specific diagnostic, prognostic, or treatment group. Several univariate measures, such as absolute and relative power, spectral ratios, phase, coherence, and symmetry may be linearly combined to form multivariate measures.[40] In doing so the measures are found to be complementary; they are additive for the detection of abnormality but yield different topographic distributions.[41] Multivariate techniques have the decided advantage of assessing the relative contributions of multiple univariate qEEG measures, thereby reducing the likelihood that important information will be overlooked.[19] This approach not only yields information about the disorder itself, but also in principle can be useful for guiding the clinical care of individual patients. For example, Prichep et al.[42] used multivariate qEEG to classify a mixed group of 54 unipolar and 23 bipolar depression patients into their correct diagnostic groups, achieving classification accuracies of 91% and 83%, respectively. To the extent that mania may ensue in a bipolar patient being treated as a unipolar patient, this might be important information for a clinician to have.

Technology Demonstrations V. Clinical Tools The use of quantitative statistical procedures for EEG abnormality detection, particularly when employed on a post hoc basis to call attention to features that might have been overlooked by the electroencephalographer during an initial visual reading, is supported by a convincing literature (discussed below). The application of such procedures to assist in clinical diagnosis by classifying a patient into the "best fit" multivariate category is less well supported by the peer reviewed literature, but a reasonable case can be made for its cautious use. Unfortunately, qEEG proponents, such as Hughes and John,[7] go beyond these modest boundaries and cite studies using techniques, such as cluster analysis that have no direct clinical application. Though cluster analysis is an important research tool and may lead to the development of clinically useful discriminants, the uncritical mixing of such studies with the more conservative citations undermines the credibility of their argument.

Another aspect of this general problem is that many qEEG studies in the literature bearing upon psychiatric problems use idiosyncratic methods of data recording and analysis involving ad hoc healthy and clinical normative groups. Such idiosyncratic research methods are of little help to the clinical psychiatrist who needs a standardized laboratory test. This problem is discussed in more depth later.

Commercial Interests Nuwer[10] questions the veracity of reports published by authors having commercial interests in qEEG systems. However, there appears to be a wide range of professionalism among authors with commercial interests, paralleling the professionalism among academic authors. Both groups profit from their endeavors, whether through promotion/tenure/salary/consulting fees or through patent royalties/corporate profits. The academic who hires himself out as an expert witness testifying against qEEG has little to distinguish himself from the commercially involved researcher promoting the technique. And although it is unfortunate that qEEG requires very large databases that are available only as commercial products, it would be difficult

to name a medical test that does not involve a commercial vendor. Scientific quality is where one finds it and the gold standard must remain articles, particularly independent replications, published in peer-reviewed journals.

A more troubling aspect of the commercial interest problem is the advertising by individual physicians. For example, in the advertising material for a recent "antiaging" seminar for clinicians in Las Vegas, a well known physician claimed that by using his "Brain Code, based on four electrical signals" derived from brain electrical activity mapping performed on a laptop computer in any doctor's office, the attendees could treat "any brain disease" including dopaminergic brain dysfunctions (Parkinson's, depression, dysthymic [sic], narcolepsy, chronic fatigue syndrome), acetyl-cholinergic [sic] brain dysfunction (Alzheimer's, memory loss, dyslexia, attention deficit disorder[ADD], cognitive disorder, mild learning disability), gamma aminobutyric acid (GABA)-dominant brain disorder (anxiety, manic depression, headaches, migraine headaches, chronic pain), and serotonin brain disorder (social phobias, insomnia, dysthymia, mixed anxiety states, somatization, irritable bowel syndrome, fibromyalgia). But as disturbing as such claims may be in this era of evidence-based medicine, the behavior of individual physicians cannot reasonably be used as a criterion for the acceptance or rejection of a laboratory procedure. The facts must speak for themselves.

Resurrecting Moot Points Many of the once-valid criticisms of qEEG have been addressed but continue to be raised by those opposing acceptance of the technique. Examples of such dead horse flogging[43,44] include standard EEG artifacts, recording errors, patient characteristics, and misapplication of techniques. It is certainly true that any of these can bias qEEG (especially multivariate) results in ways difficult to detect. Closely related are issues of technical competence and lab certification. However, these are essentially training and regulatory issues and have been dealt with through minimum practice standards, such as those proposed by Duffy et al.[9]

qEEG has been criticized for employing too many statistical tests,[43–45] thereby generating spurious statistical significance. This continues to be true of some systems using univariate comparisons and SPM to call the electroencephalographer's attention to possibly important features as discussed above. Duffy et al.[9,46] recommends replicating each clinical test and accepting as true abnormalities that replicate. The more capable qEEG systems employing multivariate comparisons use Principal Components Analysis (PCA) to reduce the large number of variables to a much smaller set of uncorrelated factors representing the intrinsic dimensionality of the data set. Either of these procedures answers the "too many statistical tests" criticism.

Another criticism[10] is that many qEEG findings are not replicated. The validity of this criticism may be judged by the reader (Tables 1 and 2). Caution should be exercised, however, in determining what constitutes a replication. Studies using discriminant analysis generally form the discriminant function from the first subject sample and test its accuracy on a second sample. In such cases the third sample would constitute the first true replication. However, the classification ability of the discriminant is assessed as early as the first sample, so has become common practice in the qEEG literature to refer to the second sample as the first replication, and this terminology has been incorporated into the present paper.

Yet another aspect of the problem is qEEG's differential sensitivity. It may be sensitive to statistically significant but clinically trivial normal variants, such as an absence of a posterior resting rhythm, while being insensitive to clinically important patterns, such as fast transient epileptiform spike activity and slower triphasic waves.[41,47] Closely related is the criticism that computers cannot diagnose disorders.[43,44] To overcome these limitations the electroencephalographer's trained eye is necessary. But qEEG does not remove the expert physician from the loop. For at least the past decade there has been universal agreement that the indispensable initial step in qEEG analysis is the "gold standard" of a clinical (visual) reading of the raw EEG by a trained electroencephalographer. qEEG is a post-hoc supplementary and complementary technique of data analysis that is specifically not intended to function as a standalone diagnostic instrument.

Standards of Evidence The argument has been made that levels of specificity found in qEEG studies are often higher than those found in routinely used clinical tests, such as mammograms, cervical screenings, or CT or SPECT brain scans.[7,48,49] The unfortunate marketing history of qEEG in the 1980s has led to a situation today in which extraordinary evidence is required for its acceptance and endorsement by professional societies. Even

TABLE 1. Diagnostic Classification Analysis

Ahn et al.[61] (1980) qEEG Classification by Heuristic Procedures

Actual Group	N	Classification (%) as* Healthy	Abnormal	PPV**	NPV**
Healthy US	306	90–96	10–4		
Healthy Barbados	91	93–98	7–2		
Neurologic	474	42–52	58–48	0.90–0.95	0.59–0.54
Learning Disabled	143	43–54	57–46	0.73–0.85	0.82–0.79
Specific Learning Disabilities	163	46–53	54–47	0.74–0.78	0.79–0.77

*Classifications based on: $p<0.05$–$p<0.01$ significance of individual scores
** PPV (positive predictive value) and NPV (negative predictive value) are shown only with regard to healthy comparison subjects

Besthorn et al.[93] (1994) qEEG Classification

Actual Group	N	Classification (%) as Alzheimer's Disease (AD)	Healthy	PPV	NPV
Alzheimer's Disease (AD)	50	72	28	0.82	0.71
Healthy	42	19	81	n/a	n/a

Brenner et al.[96] (1986) qEEG Classification

Actual Group	N	Classification (%) as AD	Healthy	PPV	NPV
AD	35	66	34	0.85	0.83
Healthy	61	7	93		
AD (MMSE*<22)	24	79	21	0.83	0.92
Healthy	61	7	93		
AD (MMSE>22)	11	36	64	0.50	0.89
Healthy	61	7	93		
		AD	Depressed		
AD	35	49	51		
Depressed	23	0	100		
AD (MMSE<22)	24	58	42		
Depressed	23	0	100		
AD (MMSE>22	11	27	73		
Depressed	23	0	100		

*MMSE = Mini-Mental State Examination

Chabot and Serfontein[70] (1996) qEEG Classification

Actual Group	N	Classification (%) as* Healthy	Attention Problems	PPV	NPV
Healthy	310	95/88	5/12		
Attention Problems, normal IQ	319	7/6	93/94	0.95/.89	0.93/0.93
Attention Problems, low IQ	88	5	95	0.91	0.97
		Normal IQ	Low IQ		
Normal IQ, attention problems	319	70/60	30/40		
Low IQ, attention problems	88	29/41	71/59		
		ADD	ADHD		
ADD	165	70/66	30/34		
ADHD	179	29/33	71/67		

*Initial discriminant/independent replication (split half)
ADD = Attention deficit disorder
ADHD = Attention deficit hyperactivity disorder

Chabot et al. (1996) qEEG Classification

Actual Group	N	Classification (%) as* Healthy	ADD/ADHD	Learning Disabled (LD)
3-way discriminant				
Healthy	310	76/66	7/3	17/32
ADD/ADHD	319	8/14	89/81	3/5
Learning Disabled (LD)	115	31/30	0/9	69/61
Slow Learning Disorder (SLD)	88	9	82	9

Actual Group	N	Classification (%) as* ADD/ADHD	LD
2-way discriminant			
ADD/ADHD	319**	93/97	7/4
LD	115**	10/16	90/84
SLD	127	13	87
ADD/ADHD, low IQ	88	97	3

LD subjects had learning disorders in 2 or more areas and IQs 65–84
SLD subjects had learning disorders in one area and normal IQs
*Initial discriminant/independent replication (split half)
**Initial discriminat/independent replicatoin (assumed split half
ADD/ADHD = Attention deficit disorder/attention deficit hyperactivity disorder

Coburn et al. (1990) qEEG Classification by % theta

Actual Group	N	Classification (%) as* Mild AD	Healthy Comparison	PPV	NPV
Mild AD	38/21	24/57	76/43	1.0/1.0	0.59/0.71
Healthy	41/22	0/0	100/000		

*Initial sample/independent replication

Coburn et al.[38] (2003) qEEG Classification by % theta

Actual Group	N	Classification (%) as* Healthy	AD	PPV	NPV
Healthy	60	53/55	47/45		
AD	45	20/24	80/76	0.56/0.56	0.78/0.75

*Classification based on Global Average Reference/Partial Average Reference data

Coutin-Churchman et al.[127] (2003) qEEG Classification

Actual Group	N	Classification (%) as* Patients	Healthy	PPV	NPV
Patients	340	84	16	0.84	0.88
Healthy	67	88	12		

Deslandes et al.[106] (2004) qEEG Multivariate Classification

		Classification (%) as*	
Actual Group	N	Dementia	Depression
Dementia	73	92/(92)	8/(8)
Depression	52	10/(12)	90/(88)

*Initial sample/(jackknife replication)

Duffy et al.[92] (1984) qEEG Classification

		Classification (%) as			
Actual Group	N	Demented	Healthy	PPV	NPV
Senile Dementia	10	90	10	0.90	0.90
Elderly Healthy	10	10	90		
Pre-senile Dementia	9	89	11	1.0	0.94
Young Healthy	15	0	100		

Huang et al.[67] (2000) qEEG Dipole Classification

Actual Group		N		Classification (%) as				
I	II	I	II	I	II	Overall	PPV	NPV
AD	Healthy	38	24	90	75	84	0.85	0.82
AD	MCI	38	31	87	68	78		
PMCI	SMCI	14	17	71	82	77		
qEEG absolute power classification								
Actual Group		N		Classification (%) as				
I	II	I	II	I	II	Overall	PPV	NPV
AD	Healthy	38	24	79	33	61	0.65	0.50
AD	MCI	38	31	95	42	71		
PMCI	SMCI	14	17	86	53	68		
qEEG relative power classification								
Actual Group		N		Classification (%) as				
I	II	I	II	I	II	Overall	PPV	NPV
AD	Healthy	38	24	87	83	85	0.89	0.80
AD	MCI	38	31	79	68	74		
PMCI	SMCI	14	17	79	94	87		
qEEG relative power classification								

MCI = mild cognitive impairment
PMCI = MCI patients progressing to AD
SMCI = MCI patients remaining stable

Jacobson et al.[104] (1993) EEG and qEEG classification

Actual Group	N	Classification (%) as Abnormal EEG	Patient	Healthy	PPV	NPV
Classification (%) by relative proportion of alpha using qEEG						
Patients (all)	27	25	93	7	0.96	0.75
Healthy	7	1	14	86		

Actual Group	N	Classification (%) as Abnormal EEG	Delirium	Dementia
Classification (%) by relative proportion of delta using qEEG				
Delirium	18	11	61	39
Dementia	9	5	44	56

Actual Group	N	Classification (%) as Abnormal EEG	Delirium	Dementia
Classification (%) by amount of theta using conventional EEG				
Delirium	18	17	94	6
Dementia	9	7	22	78

John et al.[15] (1983) qEEG classification

Actual Group	N	Classification (%) as Healthy	Mild	Mod/Sev
Healthy	89	79/(74)	17/(20)	4/(6)
Mild Dementia	50	18/(26)	60/(48)	22/(26)
Moderate/Severe Dementia	47	17/(17)	19/(23)	64/(60)

*Initial discriminant/(Jackknife replication)

Actual Group	N	Classification (%) as Healthy	Alc	Dep	Mild	Mod
Healthy	80	71	9	8	11	1
Alcoholic	31	13	74	7	3	3
Depressive	31	19	7	61	10	3
Mild Dementia	44	7	2	9	64	18
Moderate/Severe Dementia	30	10	0	10	17	63

Actual Group	N	Classification (%) as Healthy	Abnormal	PPV	NPV
qEEG classification using narrow discriminant					
Healthy 1	153	80	20		
Learning Disabled 1	89	28	72	0.67	0.83
Independent Replication					
Healthy 2	153	87	13		
Learning Disabled 2	50	35	65	0.62	0.88
Learning Disabled 3	71	35	65	0.70	0.84
Neurologic	317	29	70	0.92	0.58
Specific Learning Disabilities	159	54	49	0.78	0.61
Epileptic	85	21	79	0.77	0.88

Actual Group	N	Classification (%) as Healthy	Classification (%) as Abnormal	PPV	NPV
qEEG classification using wider discriminant					
Healthy 1	153	82	18		
Learning Disabled 1	69	39	61	0.60	
Neurologic 1	138	23	77	0.79	
Independent Replication					
Healthy 2	153	89	11		
Learning Disabled 2	69	39	61	0.71	0.83
Neurologic 2	138	30	70	0.85	0.77
Specific Learning Disabilities	159	51	49	0.82	0.63
Epileptic	178	28	72	0.88	0.73
Mental Retardation	46	28	72	0.66	0.91
Dialysis	35	6	94	0.66	0.99
Turner's Syndrome	17	24	77	0.66	0.97
Leukemia	15	47	53	0.32	0.95
Cerebral Palsy	10	30	70	0.29	0.98

Learning Disabled: No known neurological disease; 64<IQ<85; language and/or arithmetic standard scores <90
Neurologic: Neurologically at risk; including cerebral palsy, severe mental retardation, phenylketonuria, renal disease, epilepsy, severe emotional disturbances, and learning disabilities
Specific Learning Disabilities: No known neurological disease; IQ >85; language and/or arithmetic standard scores <90

John et al.[19] (1988) qEEG Classification

Actual Group	N	Classification (%) as* Healthy	Dep	Dem	Alc
Healthy	60/60	77/75	11/10	5/12	7/3
Primary Depression	69/34	9/6	72/85	9/3	10/6
Dementia	63/62	9/13	6/8	79/77	6/2
Alcoholism	20/10	20/5	0/5	0/0	80/80

Group	N	Classification (%) as* Unipolar	Bipolar
Unipolar	34/34	85/85	15/15
Bipolar	18/17	15/13	85/87

*Initial discriminant/independent replication

John and Prichep[20] (1993) qEEG Classification

Actual Group		N		Mean Discriminant Accuracy (%)*			
I	II	I	II	I	II	PPV	NPV
Two-Group Discriminants							
Healthy	Depressed	95	111	88/86	83/93	0.89	0.82
Unipolar	Bipolar	65	32	84/87	88/9		
Healthy	Schizophrenia	149	57	96/99	90/82	0.98	0.94
Depression	Schizophrenia	103	46	84/88	84/85**		
Healthy	Alcoholism	120	30	95/95	75/90	0.82	0.97
Healthy	Learning Disabled	158	175	89/79	72/71	0.79	0.71

Actual Group	N	Mean Discriminant Accuracy (%)*		
Multi-Group Discriminants				
Healthy v. Dep v. Dem	85/87/125	84/85	84/80	84/71
Vasc. Dem v. Nonvasc. Dem	91/13	94/82	92/(85)	

Dep = Depression; Dem = Dementia; Vasc Dem = Vascular Dementia; Nonvasc Dem = Nonvascular Dementia; Alc = Alcoholism
*Initial discriminant/independent replication; **Medicated group used for replication

Knott et al.[39] (1996) qEEG Classification

		Classification (%) as			
Actual Group	N	Panic Disorder	Healthy	PPV	NPV
qEEG classification by absolute amplitude					
Panic Disorder	34	71	29	0.60	0.62
Healthy	19	16	84		
		Classification (%) as			
Actual Group	N	Panic Disorder	Healthy	PPV	NPV
qEEG classification by relative amplitude					
Panic Disorder	34	65	35	0.85	0.56
Healthy	19	21	79		

Leuchter et al. (1987) qEEG Classification

		Classification (%) as			
Actual Group	N	Demented	Healthy	PPV	NPV
Demented*	18	83	17	1.0	0.67
Healthy	6	0	100		

*12 with dementia of Alzheimer's type + 6 with mild depressive disorder

Leuchter et al.[39] (1993) qEEG Classification Based on Power

Actual Group		N		Overall Correct Classification (%)
I	II	I	II	
Two-group discriminants				
DAT	Healthy	49	38	77
MID	Healthy	29	38	81
DAT	MID	49	29	69(76)*

*Classification based on coherence from Leuchter et al. (1992) in parentheses
DAT = dementia of Alzheimer's type; MID = mild depressive disorder

Lindau et al.[78] (2003) qEEG and Neuropsych Classification

Actual Group		N		Overall Correct Classification (%)		
				qEEG	Neuropsych	Combined
I	II	I	II			
FTD	Healthy	19	19	79		
AD	Healthy	16	19	80		
FTD	AD	19	16	71	80	93

FTD = formal thought disorder

Lubar et al.[62] (1985) qEEG Classification

		Classification (%) as			
Actual Group	N	Learning disabled	Healthy	PPV	NPV
LD	69	79	21	0.90	0.67
Healthy	34	19	81		

Mann et al.[68] (1992) qEEG Classification

Actual Group	N	Classification (%) as ADD	Classification (%) as Healthy	PPV	NPV
ADD	25	80	20	0.74	0.80
Healthy	27	26	74		

Mody et al. (1991) EEG and qEEG Classification

Actual Group	N	Classification (%) as AD	Classification (%) as Healthy	PPV	NPV
qEEG analysis					
AD	50*	98	2	1.0	0.98
Healthy	46	0	100		
Visual analysis of defined abnormality					
AD		16	84	1.0	0.52
Healthy		0	100		
Visual analysis of overall abnormality					
AD		98	2		
Healthy		?	?		

*5 histologically confirmed
? = Data not provided

Monastra et al.[69] (1999) qEEG Classification

Actual Group	N	Classification (%) as ADHD	Classification (%) as Healthy	PPV	NPV
ADHD	397	86	14	0.99	0.60
Healthy	85	2	98		

O'Connor et al. (1979) qEEG Classification

Actual Group	N	Classification (%) as Demented	Classification (%) as Depressed
Demented	16	88	12
Depressed	8	0	100
		Senile Dementia	Arteriosclerosis
Senile Dementia	8	88	12
Arteriosclerosis	8	0	100

Prichep and John[115] (1986) qEEG Classification

Actual Group	N	Classification (%) as* Healthy	Dep	Dem	Alc
Healthy	93	84/(76)	7/(11)	5/(5)	4/(8)
Depressed	51	13/(15)	73/(65)	10/(14)	4/(6)
Dementia	47	13/(13)	13/(13)	73/(73)	0
Alcoholism	31	8/(12)	8/(14)	10/(10)	74/(64)

*Initial discriminant/(jackknife replication)

Prichep and John[115] (1986) qEEG Classification *continued*

Actual Group	N	Classification (%) as*		PPV	NPV
		Healthy	Depressed		
Healthy	93	89/(87)	11/(13)		
Depressed (primary + secondary)	70	17/(19)	83/(81)	0.85/(0.83)	0.87/(0.86)
		Unipolar	Bipolar		
Unipolar	31	87/(87)	13/(13)		
Bipolar	20	10/(15)	90/(85)		

*Initial discriminant/(Jackknife replication)

Prichep et al.[42] (1990) qEEG and QEP Classification

Actual Group	N	Classification (%) as*	
		Unipolar	Bipolar
Classification using qEEG variables only			
Unipolar	54/54	91/76	9/24
Bipolar	23/20	17/25	83/75
		Classification (%) as*	
Actual Group	N	Unipolar	Bipolar
Classification using qEEG and QEP variables			
Unipolar	48/45	98/76	2/24
Bipolar	21/17	9/18	91/82

*Initial discriminant/Independent replication

Prinz and Vitiello[88] (1989) EEG Classification

Actual Group	N	Classification (%) as		PPV	NPV
		Early AD	Healthy		
Visually analyzed EEG (alpha slowing)					
Early AD	41	71	29	0.76	0.77
Healthy	50	18	82		

*(possible + probable)

Pritchard et al.[103] (1994) qEEG Classification by Different Methods

Actual Group	N	Classification (%) as*		PPV	NPV
		Healthy	AD		
Healthy	25	80/96	20/4		
AD	14	28/14	72/86	0.67/0.92	0.83/0.92

*Classification based on Nearest Neighbor Discriminant/Neural Net analyses

Robinson et al.[89] (1994) EEG Classification (Adapted, Combined Controls)

		Classification (%) as			
Actual Group	N	Healthy	Abnormal	PPV	NPV
Healthy	56	64	36		
AD	86	13	87	0.79	0.77
AD + MID	17	24	76	0.39	0.90

Schreiter-Gasser et al.[94] (1993) qEEG Classification

		Classification (%) as			
Actual Group	N	AD	Healthy	PPV	NPV
AD	15	93	7	1.0	0.94
Healthy	15	0	100		

Sloan et al. (1995) EEG Classification

		Classification (%) as*			
Actual Group	N	AD	MID	MDD	
AD	43	77	12	12	
MID	25	0	76	24	
MDD	29	7	14	79	
		Healthy	Abnormal	PPV	NPV
MDD	29	79	01		
AD	43	12	88	0.86	0.82
MID	25	24	76	0.76	0.79

*MDD assumed to have visually normal EEG

Streletz et al.[100] (1990) qEEG Classification

		Classification (%) as			
Actual Group	N	Demented	Healthy	PPV	NPV
DAT	25	68	32	0.94	0.53
Healthy*	10	10	90		
PD	10	70	30	0.88	0.75
Healthy*	10	10	90		

*Separate healthy groups, each age appropriate

Yener et al.[90] (1996) EEG and qEEG Classification

Actual Group		N		Classification (%) as*			
I	II	I	II	I	II	PPV	NPV
qEEG classification							
AD	Healthy	26	27	77/(82)	89/(73)	0.87/(0.75)	0.80/(0.80)
FTD	Healthy	13	27	85/(77)	100/(100)	1.0/(1.0)	0.93/(0.90)
AD	FTD	26	13	85/(81)	85/(85)		

		Classification (%) as*			
Actual Group	N	Healthy	Abnormal	PPV	NPV
Healthy	27	93	7		
AD	26	15	85	0.92	0.86
FTD	13	31	69	0.82	0.86

*Initial discriminant/(Jackknife replication)

FDA findings of safety and efficacy do not appear to be sufficient. The opponent camp championed by Nuwer maintains that additional evidence is needed, while the proponent camp championed by John counters that existing evidence is overlooked or misinterpreted. It is possible that clinical turf issues may play a role in this dispute. However, it has yet to be shown that any qEEG system available to the working clinical psychiatrist meets the methodological standards for diagnostic tests (spectrum composition, analysis of pertinent subgroups, avoidance of workup bias, avoidance of review bias, precision of results for test accuracy, presentation of indeterminate test results, test reproducibility) enumerated by Reid et al.[50]

Information Availability A major problem faced by the psychiatrist wishing to assess the practical clinical usefulness of commercial qEEG systems is that information about most systems' capabilities is extremely difficult to obtain. The FDA has in the past placed severe restrictions on the information available to potential users, even forbidding a listing of the specific analyses available, and the ludicrous situation has arisen wherein, even after purchasing one major system, the buyer finds no such listing in the user manual. The situation may be changing since the most recently approved system is much better described. Lawsuits between commercial vendors similarly constrain the information they make available.

Applications to Specific Disorders

When reading the following sections, an important point must be kept in mind. Each section contains comparisons between standard, visually analyzed EEG and qEEG. Preceding parts of this article have stressed repeatedly that the indispensable first step in qEEG analysis is a standard visual reading of the raw EEG by a qualified electroencephalographer, and that qEEG is used responsibly only as a supplementary and complementary technique for post hoc analysis, serving to draw the physician's attention to aspects of the original EEG that may have been overlooked. It is always the physician who performs the diagnosis or makes other relevant clinical decisions, not the machine. However, in the sections to follow, these fundamental dicta are consistently violated in order to assess the ability of the

TABLE 2. Drug Response Prediction Analyses

Chabot et al.[71] (1996) qEEG Classification

		Classification (%) as*	
Actual Group	N	DEX	MPH
Dextroamphetamine Responders	65	76/69	24/31
Methylphenidate Responders	81	24/32	76/68

*Initial discriminant/independent replication (split half)

Chabot et al.[53] (1999) qEEG Classification

		Classification (%) as*	
Actual Group	N	Responders	Nonresponders
Stimulant Responders	76	83/83	17/17
Stimulant Nonresponders	54	12/12	88/88

*Initial discriminant/independent replication (split half)

Galderisi et al.[126] (1994) qEEG Classification

		Classification (%) as	
Actual Group	N	Responders	Nonresponders
Antipsychotic Responders	18	94	6
Antipsychotic Nonresponders	10	20	80

Hansen et al.[26] (2003) qEEG Cluster Analysis

		Classification (%) as	
SSRI Response	N	Cluster 1	Cluster 2
Responder	18	50	94
Nonresponder	2	50	